

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Nelson, Edward L.
Nelson, Peter J.
- (ii) TITLE OF INVENTION: NOVEL VECTOR FOR
POLYNUCLEOTIDE VACCINES
- (iii) NUMBER OF SEQUENCES: 29
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 - (B) STREET: 345 PARK AVENUE
 - (C) CITY: NEW YORK
 - (D) STATE: NEW YORK
 - (E) COUNTRY: USA
 - (F) ZIP: 10154
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
 - (B) COMPUTER: IBM PC COMPATIBLE
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: WORDPERFECT 5.1
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/US97/14306
 - (B) FILING DATE: 14-AUG-1997
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US60/023931
 - (B) FILING DATE: 14-AUG-1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: KATHRYN M. BROWN
 - (B) REGISTRATION NUMBER: 34556
 - (C) REFERENCE/DOCKET NUMBER: 2026-4236US1
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 758-4800
 - (B) TELEFAX: (212) 751-6849
 - (C) TELEX: 421792

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 453 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:1:

| | | | | |
|------------|------------|-------------|------------|-----|
| GGCCGCGTTG | CTGGCGTTTT | TCCATAGGCT | CCGCCCCCCT | 40 |
| GACGAGCATC | ACAAAAATCG | ACGCTCAAGT | CAGAGGTGGC | 80 |
| GAAACCCGAC | AGGACTATAA | AGATAACCAGG | CGTTTCCCCC | 120 |
| TGGAAGCTCC | CTCGTGCGCT | CTCCTGTTCC | GACCCTGCCG | 160 |
| CTTACCGGAT | ACCTCTCCGC | CTTTCTCCCT | TCGGGAAGCG | 200 |
| TGGCGCTTTC | TCAATGCTCA | CGCTGTAGGT | ATCTCAGTTC | 240 |
| GGTGTAGGTC | GTTCGCTCCA | AGCTGGGCTG | TGTGCACGAA | 280 |
| CCCCCGTTT | AGCCCGACCG | CTGCGCCTTA | TCCGGTAACT | 320 |
| ATCGTCTTGA | GTCCAACCCG | GTAAGACACG | ACTTATCGCC | 360 |
| ACTGGCAGCA | GCCACTGGTA | ACAGGATTAG | CAGAGCGAGG | 400 |
| TATGTAGGCG | GTGCTACAGA | GTTCTTGAAG | TGGTGGCCTA | 440 |
| ACTACGGCTA | CAC | | | 453 |

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 453 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:2:

| | | | | |
|------------|------------|------------|------------|-----|
| GTGTAGCCGT | AGTTAGGCCA | CCACTTCAAG | AACTCTGTAG | 40 |
| CACCGCCTAC | ATACCTCGCT | CTGCTAATCC | TGTTACCAGT | 80 |
| GGCTGCTGCC | AGTGCGGATA | AGTCGTGTCT | TACCGGGTTG | 120 |
| GACTCAAGAC | GATAGTTACC | GGATAAGGCG | CAGCGGTCGG | 160 |
| GCTGAACGGG | GGGTTCGTGC | ACACAGCCCA | GCTTGGAGCG | 200 |
| AACGACCTAC | ACCGAACTGA | GATACCTACA | CCGTGAGCAT | 240 |
| TGAGAAAGCG | CCACGCTTCC | CGAAGGGAGA | AAGGCGGACA | 280 |
| GGTATCCGGT | AAGCGGCAGG | GTCGGAACAG | GAGAGCGCAC | 320 |
| GAGGGAGCTT | CCAGGGGGAA | ACGCCTGGTA | TCTTTATAGT | 360 |
| CCTGTCGGGT | TTCGCCACCT | CTGACTTGAG | CGTCGATTTT | 400 |
| TGTGATGCTC | GTCAGGGGGG | CGGAGCCTAT | GGAAAAACGC | 440 |
| CAGCAACGCG | GCC | | | 453 |

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 209 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:3:

| | | | | |
|------------|------------|------------|------------|-----|
| GAATTCTTTC | GGACTTTTGA | AAGTGATGGT | GGTGGCCGAA | 40 |
| GGATTCGAAC | CTTCGAAGTC | GATGACGGCA | GATTTAGAGT | 80 |
| CTGCTCCCTT | TGGCCGCTCG | GGAACCCAC | CACGGGTAAT | 120 |
| GCTTTTACTG | GCCTGCTCCC | TTATCGGGAA | GCGGGGCGCA | 160 |
| TCATATCAAA | TGACGCGCCG | CTGTAAAGTG | TTACGTTGAG | 200 |
| AAAGAATTC | | | | 209 |

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 209 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:4:

| | | | | |
|------------|------------|------------|------------|-----|
| GAATTCTTTC | TCAACGTAAC | ACTTTACAGC | GGCGCGTCAT | 40 |
| TTGATATGAT | GCGCCCCGCT | TCCCGATAAG | GGAGCAGGCC | 80 |
| AGTAAAAGCA | TTACCCGTGG | TGGGGTTCCC | GAGCGGCCAA | 120 |
| AGGGAGCAGA | CTCTAAATCT | GCCGTCATCG | ACTTCGAAGG | 160 |
| TTCGAATCCT | TCCCCACCA | CCATCACTTT | CAAAAGTCCG | 200 |
| AAAGAATTC | | | | 209 |

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:5:

AATAAA

6

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:6:

ATTAAA

6

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:7:

AGTAAA

6

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:8:

AAGAAC

6

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:9:

AATACA

6

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 227 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:10:

| | | | | |
|------------|-------------|------------|------------|-----|
| GCCTTAAGGG | CCATATGGTG | AGTGGATCCC | TTGACCCCAG | 40 |
| GCGGGGATGG | GGAGACCTGT | AGTCAGAGCC | CCCGGGCAGC | 80 |
| ACAGGCCAAT | GCCCCGTCCTT | CCCCTGCAGG | ATGAGTAGTG | 120 |
| AGTGCCTCTC | CTGGCCCTGG | AAGTTGCCAC | TCCAGTGCCC | 160 |
| ACCAGCCTTG | TCCTAATAAA | ATTAAGTTGC | ATCATTTTGT | 200 |
| CTGACTAGGT | GTCCTCTATA | ATATTAT | | 227 |

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 227 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:11:

| | | | | |
|------------|------------|------------|------------|-----|
| ATAATATTAT | AGAGGACACC | TAGTCAGAAC | AAATGATGCA | 40 |
| ACTTAATTTT | ATTAGGACAA | GGCTGGTGGG | CACTGGAGTG | 80 |
| GCAACTTCCA | GGGCCAGGAG | AGGCACTCAC | TACTCATCCT | 120 |
| GCAGGGGAAG | GACGGGCATT | GGCCTGTGCT | GCCCGGGGGC | 160 |
| TCTGACTACA | GGTCTCCCCC | ATCCCCGCCT | GGGGTCAAGG | 200 |
| CATCCACTCA | CCATATGGCC | CTTAAGG | | 227 |

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 252 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:12:

| | | | | |
|------------|------------|------------|------------|-----|
| CCTCGGTACC | TGCCATGGCG | CGGATTCTTT | ATCACTGATA | 40 |
| AGTTGGTGGG | CATATTATGT | TTATCAGTGA | TAAAGTGTCA | 80 |
| AGCATGACAA | AGTTGCAGCC | GAATACAGTG | ATCCGTGCCG | 120 |
| GCCCTGGACT | GTTGAACGAG | GTCGGCGTAG | ACGGTCTGAC | 160 |
| GACACGCAAA | CTGGCGGAAC | GGTTGGGGGT | GCAGCAGCCG | 200 |
| GCGCTTTACT | GGCACTTCAG | GAACAAGCGG | GCGCCTTAAG | 240 |
| GGCCATATGC | CG | | | 252 |

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:13:

CCTCGGTACC TGCCACCATG GCGCGGATTC TTTAT

35

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:14:

CGGCATATGG CCTTAAGGCG CCCGCTTGTT CCTGAAGT

38

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 228 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:15:

GCCTTAAGGG CCATATGGTG AGTGGATGCC TTGACCCAG
GCGGGGATGG GGGAGACCTG TAGTCAGAGC CCCCGGGCAG

40

80

CACAGGCCAA TGCCCGTCCT TCCCCTGCAG GATGAGTAGT
GAGTGCCTCT CCTGGCCCTG GAAGTTGCCA CTCCAGTGCC
CACCAGCCTT GTCCTAATAA AATTAAGTTG CATCATTTTG
TCTGACTAGG TGTCTCTAT AATATTAT

120

160

200

228

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1425 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:16:

| | | | | |
|------------|-------------|------------|------------|------|
| TGCCATGGCG | CGGATTCTTT | ATCACTGATA | AGTTGGTGG | 40 |
| CATATTATGT | TTATCAGTGA | TAAAGTGTCA | AGCATGACAA | 80 |
| AGTTGCAGCC | GAATACAGTG | ATCCGTGCCG | GCCCTGGACT | 120 |
| GTTGAACGAG | GTCGGCGTAG | ACGGTCTGAC | GACACGCAAA | 160 |
| CTGGCGGAAC | GGTTGGGGGT | GCAGCAGCCG | GCGCTTTACT | 200 |
| GGCACTTCAG | GAACAAGCGG | GCGCCTTAAG | GGCCATATGG | 240 |
| TGAGTGGATG | CCTTGACCCC | AGGCGGGGAT | GGGGGAGACC | 280 |
| TGTAGTCAGA | GCCCCCGGGC | AGCACAGGCC | AATGCCCGTC | 320 |
| CTTCCCCTGC | AGTGAGTAGT | GACTGCCCGG | GTGGGATCCC | 360 |
| TGTGACCCCT | CCCCAGTGCC | TCTCCTGGCC | CTGGAAGTTG | 400 |
| CCACTCCAGT | GCCCCACCAGC | CTTGTCTTAA | TAAAATTAAG | 440 |
| TTGCATCATT | TTGTCTGACT | AGGTGTCCTC | TATAATATTA | 480 |
| TAAGCTTGAT | ATCGAATTCT | TTCTCAACGT | AACACTTTAC | 520 |
| AGCGGCGCGT | CATTTGATAT | GATGCGCCCC | GCTTCCCGAT | 560 |
| AAGGGAGCAG | GCCAGTAAAA | GCATTACCCG | TGGTGGGGTT | 600 |
| CCCGAGCGGC | CAAAGGGAGC | AGACTCTAAA | TCTGCCGTCA | 640 |
| TCGACTTCGA | AGGTTCTGAAT | CCTTCCCCCA | CCACCATCAC | 680 |
| TTTCAAAAGT | CCGAAAGAAT | TCCTGCAGCC | CGTGTAGCCG | 720 |
| TAGTTAGGCC | ACCACTTCAA | GAACTCTGTA | GCACCGCCTA | 760 |
| CATACCTCGC | TCTGCTAATC | CTGTTACCAG | TGGCTGCTGC | 800 |
| CAGTGGCGAT | AAGTCGTGTC | TTACCGGGTT | GGAATCAAGA | 840 |
| CGATAGTTAC | CGGATAAGGC | GCAGCGGTCT | GGCTGAACGG | 880 |
| GGGGTTCGTG | CACACAGCCC | AGCTTGGAGC | GAACGACCTA | 920 |
| CACCGAAGTG | AGATACCTAC | AGCGTGAGCA | TTGAGAAAGC | 960 |
| GCCACGCTTC | CCGAAGGGAG | AAAGGCGGAC | AGGTATCCGG | 1000 |
| TAAGCGGCAG | GGTCGGAACA | GGAGAGCGCA | CGAGGGAGCT | 1040 |
| TCCAGGGGGA | AACGCCTGGT | ATCTTTATAG | TCCTGTCGGG | 1080 |
| TTTCGCCACC | TCTGACTTGA | GCGTCGATTT | TTGTGATGCT | 1120 |
| CGTCAGGGGG | GCGGAGCCTA | TGGAAAAACG | CCAGCAACGC | 1160 |
| GGCCGGGGGA | TCCGGAGAGC | TCACTCTAGA | TGAGAGAGCA | 1200 |
| GTGAGGGAGA | GACAGAGACT | CGAATTTCCG | GAGCTATTTC | 1240 |
| AGTTTTCTTT | TCCGTTTGTG | GCAATTTTAC | TTATGATACC | 1280 |
| GGCCAATGCT | TGGTTGCTAT | TTTGGAAACT | CCCCTTAGGG | 1320 |
| GATGCCCTTC | AACTGGCCCT | ATAAAGGGCC | AGCCTGAGCT | 1360 |
| GCAGAGGATT | CCTGCAGAGG | ATCAAGACAG | CACGTGGACC | 1400 |
| TCGCACAGCC | TCTCCACAG | GTACC | | 1425 |

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 719 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:17:

| | | | | |
|-------------|------------|------------|------------|-----|
| ATGAGCAAGG | GCGAGGAACT | GTTCACTGGC | GTGGTCCCAA | 40 |
| TTCTCGTGGA | ACTGGATGGC | GATGTGAATG | GGCACAAATT | 80 |
| TTCTGTCAGC | GGAGAGGGTG | AAGGTGATGC | CACATACGGA | 120 |
| AAGCTCACCC | TGAAATTCAT | CTGCACCACT | GGAAAGCTCC | 160 |
| CTGTGCCATG | GCCAACACTG | GTCACTACCT | TCACCTATGG | 200 |
| CGTGCAAGTGC | TTTTCAGAT | ACCCAGACCA | TATGAACGAG | 240 |
| CATGACTTTT | TCAAGAGCGC | CATGCCCGAG | GGCTATGTGC | 280 |
| AGGAGAGAAC | CATCTTTTTC | AAAGATGACG | GGAACACAA | 320 |
| GACCCGCGCT | GAAGTCAAGT | TCGAAGGTGA | CACCCTGGTG | 360 |
| AATAGAATCG | AGTTGAAGGG | CATTGACTTT | AAGGAAGATG | 400 |
| GAAACATTCT | CGGCCACAAG | CTGGAATACA | ACTATAACTC | 440 |
| CCACAATGTG | TACATCATGG | CCGACAAGCA | AAAGAATGGC | 480 |
| ATCAAGGTCA | ACTTCAAGAT | CAGACACAAC | ATTGAGGATG | 520 |
| GATCCGTGCA | GCTGGCCGAC | CATTATCAAC | AGAACACTCC | 560 |
| AATCGGCGAC | CGCCCTGTGC | TCCTCCCAGA | CAACAATTAC | 600 |
| CTGTCCACCC | AGTCTGCCCT | GTCTAAAGAT | CCCAACGAAA | 640 |
| AGAGAGACCA | CATGGTCCTG | CTGGAGTTTG | TGACCGCTGC | 680 |
| TGGGATCACA | CATGGCATGG | ACGAGCTGTA | CAAGTGAGC | 719 |

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1911 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:18:

| | | | | |
|-------------|------------|------------|-------------|------|
| TATGAGCAAG | GGCGAGGAAC | TGTTCACTGG | CGTGGTCCCA | 40 |
| ATTCTCGTGG | AACTGGATGG | CGATGTGAAT | GGGCACAAAT | 80 |
| TTTCTGTCAG | CGGAGAGGGT | GAAGGTGATG | CCACATACGG | 120 |
| AAAGCTCACC | CTGAAATTCA | TCTGCACCAC | TGGAAAGCTC | 160 |
| CCTGTGCCAT | GGCCAACACT | GGTCACTACC | TTCACCTATG | 200 |
| GCGTGCAGTG | CTTTTCCAGA | TACCCAGACC | ATATGAAGCA | 240 |
| GCATGACTTT | TTCAAGAGCG | CCATGCCCCG | GGGCTATGTG | 280 |
| CAGGAGAGAA | CCATCTTTTT | CAAAGATGAC | GGGAACCTACA | 320 |
| AGACCCGCGC | TGAAGTCAAG | TTCGAAGGTG | ACACCCTGGT | 360 |
| GAATAGAATC | GAGTTGAAGG | GCATTGACTT | TAAGGAAGAT | 400 |
| GGAAACATTC | TCGGCCACAA | GCTGGAATAC | AACTATAACT | 440 |
| CCCACAATGT | GTACATCATG | GCCGACAAGC | AAAAGAATGG | 480 |
| CATCAAGGTC | AACTTCAAGA | TCAGACACAA | CATTGAGGAT | 520 |
| GGATCCGTGC | AGCTGGCCGA | CCATTATCAA | CAGAACACTC | 560 |
| CAATCGGCGA | CGGCCCTGTG | CTCCTCCCAG | ACAACCATTA | 600 |
| CCTGTCCACC | CAGTCTGCCC | GTCTAAAGAT | CCCAACGAAA | 640 |
| AGAGAGACCA | CATGGTCCTG | CTGGAGTTTG | TGACCGCTGC | 680 |
| TGGGATACAC | CATGGCATGG | ACGAGCTGTA | CAGATGAGCC | 720 |
| ATATGGTGAG | TGGATGCCTT | GACCCCAGGC | GGGGATGGGG | 760 |
| GAGACCTGTA | GTCAGAGCCC | CCGGGCAGCA | CAGGCCAATG | 800 |
| CCCGTCCTTC | CCCTGCAGTG | AGTAGTGAAT | GCCCCGGTGG | 840 |
| GATCCCTGTG | ACCCCTCCCC | AGTGCCCTCT | CTGGCCCTGG | 880 |
| AAGTTGCCAC | TCCAGTGCCC | ACCAGCCTTG | TCCTAATAAA | 920 |
| ATTAAGTTGC | ATCATTTTGT | CTGACTAGGT | GTCTCTCTATA | 960 |
| ATATTATAAG | CTTGATATCG | AATTCTTTCT | CAACGTAACA | 1000 |
| CTTTACAGCG | GCGCGTCATT | TGATATGATG | CGCCCCGCTT | 1040 |
| CCCGATAAGG | GAGCAGGCCA | GTAAAAGCAT | TACCCGTGGT | 1080 |
| GGGGTTCCCG | AGCGGCCAAA | GGGAGCAGAC | TCTAAATCTG | 1120 |
| CCGTCATCGA | CTTCGAAGGT | TCGAATCCTT | CCCCCACCAC | 1160 |
| CATCACTTTC | AAAAGTCCGA | AAGAATTCCT | GCAGCCCGTG | 1200 |
| TAGCCGTAGT | TAGGCCACCA | CTTCAAGAAC | TCTGTAGCAC | 1240 |
| CGCCTACATA | CCTCGCTCTG | CTAATCCTGT | TACCACTGGC | 1280 |
| TGCTGCCAGT | GGCGATAAGT | CGTGTCTTAC | CGGGTTGGAC | 1320 |
| TCAAGACGAT | AGTTACCGGA | TAAGGCGCAG | CGGTCGGGCT | 1360 |
| GAACGGGGGG | TTCGTGCACA | CAGCCCAGCT | TGGAGCGAAC | 1400 |
| GACCTACACC | GAACTGAGAT | ACCTACAGCG | TGAGCATTGA | 1440 |
| GAAAGCGCCA | CGCTTCCCGA | AGGGAGAAAG | GCGGACAGGT | 1480 |
| ATCCGGTAAG | CGGCAGGGTC | GGAACAGGAG | AGCGCACGAG | 1520 |
| GGAGCTTCCA | GGGGGAAACG | CCTGGTATCT | TTATAGTCCT | 1560 |
| GTCGGGTTTC | GCCACCTCTG | ACTTGAGCGT | CGATTTTTGT | 1600 |
| GATGCTCGTC | AGGGGGGCGG | AGCCTATGGA | AAAACGCCAG | 1640 |
| CAACGCGGCC | GGGGGATCCG | GAGAGCTCAC | TCTAGATGAG | 1680 |
| AGAGCAGTGA | GGGAGAGACA | GAGACTCGAA | TTTCCGGAGC | 1720 |
| TATTTTCAGTT | TTCTTTTCCG | TTTTGTGCAA | TTTCACTTAT | 1760 |
| GATACCGGCC | AATGCTTGGT | TGCTATTTTG | GAAACTCCCC | 1800 |
| TTAGGGGATG | CCCCTCAACT | GGCCCTATAA | AGGGCCAGCC | 1840 |
| TGAGCTGCAG | AGGATTCTCT | CAGAGGATCA | AGACAGCAGC | 1880 |
| TGGACCTCGC | ACAGCCTCTC | CCACAGGTAC | C | 1911 |

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 69 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:19:

```

Pro Asp Leu Ser Tyr Met Pro Ile Trp Lys Phe Pro
 1      5      10
Asp Glu Glu Gly Ala Cys Gln Pro Cys Pro Ile Asn
      15      20
Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly
25      30      35
Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser
      40      45
Ile Ile Ser Ala Val Val Gly Ile Leu Leu Val Val
50      55      60
Val Leu Gly Val Val Phe Gly Ile Leu
      65

```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 287 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:20:

```

Pro Ala Pro Gly Ala Gly Gly Met Val His His Arg
 1      5      10
His Arg Ser Ser Ser Thr Arg Ser Gly Gly Gly Asp
      15      20
Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala
25      30      35
Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly
      40      45
Ser Asp Val Phe Asp Gly Asp Leu Gly Met Gly Ala
50      55      60
Ala Lys Gly Leu Ser Leu Pro Thr His Asp Pro Ser
      65      70
Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro
75      80
Leu Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu
85      90      95

Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro
      100      105
Asp Val Arg Pro Pro Pro Ser Pro Arg Glu Gly Pro
110      115      120

Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu
      125      130
Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val
135      140

```

| | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Lys | Asp | Val | Phe | Ala | Phe | Gly | Gly | Ala | Val | Glu |
| 145 | | | | | 150 | | | | | 155 | |
| Asn | Pro | Glu | Tyr | Leu | Thr | Pro | Gln | Gly | Thr | Cys | Ser |
| | | 160 | | | | | | 165 | | | |
| Pro | Gln | Pro | Glu | Tyr | Val | Asn | Gln | Pro | Asp | Val | Arg |
| | 170 | | | | 175 | | | | | | 180 |
| Pro | Gln | Pro | Pro | Ser | Pro | Arg | Glu | Gly | Pro | Leu | Pro |
| | | | 185 | | | | | | 190 | | |
| Ala | Ala | Arg | Pro | Ala | Gly | Ala | Thr | Leu | Glu | Arg | Pro |
| | | 195 | | | | | 200 | | | | |
| Lys | Leu | Ser | Pro | Gly | Lys | Asn | Gly | Val | Val | Lys | Asp |
| 205 | | | | | 210 | | | | | 215 | |
| Val | Phe | Ala | Phe | Gly | Gly | Ala | Val | Glu | Asn | Pro | Glu |
| | | | 220 | | | | | 225 | | | |
| Tyr | Leu | Thr | Pro | Gln | Gly | Gly | Ala | Ala | Pro | Gln | Pro |
| | 230 | | | | 235 | | | | | | 240 |
| His | Pro | Pro | Pro | Ala | Phe | Ser | Pro | Ala | Phe | Asp | Asn |
| | | | | 245 | | | | | 250 | | |
| Leu | Tyr | Tyr | Trp | Asp | Asp | Pro | Pro | Glu | Arg | Gly | Ala |
| | | 255 | | | | | 260 | | | | |
| Pro | Pro | Ser | Thr | Phe | Lys | Gly | Thr | Pro | Thr | Ala | Glu |
| 260 | | | | | 270 | | | | | 275 | |
| Asn | Pro | Glu | Tyr | Leu | Gly | Leu | Asp | Val | Pro | Val | |
| | | | 280 | | | | | 285 | | | |

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

| | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Ile | Ser | Ala | Val | Val | Gly | Ile | Leu | Leu | Val | Val |
| 1 | | | | 5 | | | | | 10 | | |
| Val | Leu | Gly | Val | Val | Phe | Gly | Ile | Leu | Ile | | |
| | | 15 | | | | | 20 | | | | |

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2125 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:22:

| | | | | |
|-------------|------------|------------|-------------|------|
| GCCACCATGG | CCCCTGACCT | CTCCTACATG | CCCATCTGGA | 40 |
| AGTTTCCAGA | TGAGGAGGGC | GCATGCCAGC | CTTGCCCCAT | 80 |
| CAACTGCACC | CACTCCTGTG | TGGACCTGGA | TGACAAGGGC | 120 |
| TGCCCCGCCG | AGCAGAGAGC | CAGCCCTCTG | ACGTCCATCA | 160 |
| TCTCTGCGGT | GGTTGGCATT | CTGCTGGTCG | TGGTCTTGGG | 200 |
| GGTGGTCTTT | GGGATCCTCA | TCAAGCGACG | GCAGCAGAAG | 240 |
| ATCACATGTC | CAGACCCTGC | CCCGGGCGCT | GGGGGCATGG | 280 |
| TCCACCACAG | GCACCGCAGC | TCATCTACCA | GGAGTGCGG | 320 |
| TGGGGACCTG | ACACTAGGGC | TGGAGCCCTC | TGAGAGGAG | 360 |
| GCCCCACGGT | CTCCACTGGC | ACCCTCCGAA | GGGGCTGGCT | 400 |
| CCGATGTATT | TGATGGTGAC | CTGGGAATGG | GGGCAGCCAA | 440 |
| GGGGCTGCAA | AGCCTCCCCA | CACATGACCC | CAGCCCTCTA | 480 |
| CAGCGGTACA | GTGAGGACCC | CACAGTACCC | CTGCCCTCTG | 520 |
| AGACTGATGG | CTACGTTGCC | CCCCTGACCT | GCAGCCCCCA | 560 |
| GCCTGAATAT | GTGAACCAGC | CAGATGTTCT | GCCCCAGCCC | 600 |
| CCTTCGCCCC | GAGAGGGCCC | TCTGCCTGCT | GCCCCGACCTG | 640 |
| CTGGTGCCAC | TCTGGAAAGG | CCCAAGACTC | TCTCCCCAGG | 680 |
| GAAGAATGGG | GTCGTCAAAG | ACGTTTTTGC | CTTTGGGGGT | 720 |
| GCCGTGGAGA | ACCCCGAGAC | TTGACACCCC | AGGGAGGAGC | 760 |
| TGCCCCCTCAG | CCCCACCCTC | CTCCTGCCTT | CAGCCCAGCC | 800 |
| TTGACAACCC | TCTATTACTG | GGACCAGGAC | CCACCAGAGC | 840 |
| GGGGGGCTCC | ACCCAGCACC | TTCAAAGGGA | CACCTACGGC | 880 |
| AGAGAACCCA | GAGTACCTGG | GTCTGGACGT | GCCAGTGTGA | 920 |
| AGCCTTAAGG | GCCATATGGT | GAGTGGATGC | CTTGACCCCA | 960 |
| GGCGGGGATG | GGGGAGACCT | GTAGTCAGAG | CCCCCGGGCA | 1000 |
| GCACAGGCCA | ATGCCCGTCC | TTCCCCTGCA | GTGAGTAGTG | 1040 |
| ACTGCCCGGG | TGGGATCCCT | GTGACCCCTC | CCCAGTGCCT | 1080 |
| CTCCTGGCCC | TGGAAGTTGC | CACTCCAGTG | CCCACCAGCC | 1120 |
| TTGTCCTAAT | AAAATTAAGT | TGCATCATTT | TGTCTGACTA | 1160 |
| GGTGTCTCT | ATAATATTAT | AAGCTTGATA | TGAATTCTT | 1200 |
| TCTCAACGTA | ACACTTTACA | GCGGCGCGTG | ATTTGATATG | 1240 |
| ATGCGCCCCG | CTTCCCGATA | AGGGAGCAGC | CCAGTAAAAG | 1280 |
| CATTACCCGT | GGTGGGGTTC | CCGAGCGGCC | AAAGGGAGCA | 1320 |
| GACTCTAAAT | CTGCCGTCAT | CGACTTCGAA | GGTTCGAATC | 1360 |
| CTTCCCCCAC | CACCATCACT | TTCAAAAGTC | CGAAAGAATT | 1400 |
| CCTGCAGCCC | GTGTAGCCGT | AGTTAGGCCA | CCACTTCAAG | 1440 |
| AACTCTGTAG | CACCGCCTAC | ATACCTCGCT | CTGCTAATCC | 1480 |
| TGTTACCAGT | GGCTGCTGCC | AGTGGCGATA | AGTCGTGTCT | 1520 |
| TACCGGGTTG | GACTCAAGAC | GATAGTTACC | GGATAAGGCG | 1560 |
| CAGCGGTCGG | GCTGAACGGG | GGGTTCTGTC | ACACAGCCCA | 1600 |
| GCTTGGAGCG | AACGACCTAC | ACCGAACTGA | GATACCTACA | 1640 |
| GCGTGAGCAT | TGAGAAAGCG | CCACGCTTCC | CGAAGGGAGA | 1680 |
| AAGGCGGACA | GGTATCCGGT | AAGCGGCAGG | GTCGGAACAG | 1720 |
| GAGAGCGCAC | GAGGGAGCTT | CCAGGGGGAA | ACGCCTGGTA | 1760 |
| TCTTTATAGT | CCTGTCGGGT | TTCGCCACCT | CTGACTTGAG | 1800 |
| CGTCGATTTT | TGTGATGCTC | GTCAGGGGGG | CGGAGCCTAT | 1840 |
| GGAAAAACGC | CAGCAACGCG | GCCGGGGGAT | CCGGAGAGCT | 1880 |
| CACTCTAGAT | GAGAGAGCAG | TGAGGGAGAG | ACAGAGACTC | 1920 |
| GAATTTCCGG | AGCTATTTCA | GTTTTCTTTT | CCGTTTTGTG | 1960 |

| | | | | |
|------------|------------|------------|------------|------|
| CAATTTCACT | TATGATACCG | GCCAATGCTT | GGTTGCTATT | 2000 |
| TTGGAAACTC | CCCTTAGGGG | ATGCCCCTCA | ACTGGCCCTA | 2040 |
| TAAAGGGCCA | GCCTGAGCTG | CAGAGGATTC | CTGCAGAGGA | 2080 |
| TCAAGACAGC | ACGTGGACCT | CGCACAGCCT | CTCCCACAGG | 2120 |
| TACCT | | | | 2125 |

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: oligonucleotide
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:23:

GTCTGCCACC ATGGCCTACT CCCCTGC

27

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: oligonucleotide
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION:SEQ ID NO:24:

TTCTTTGGTG ACCTACCTCT TCGGAATTGC CGAGTC

36

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1242 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

| | | | | |
|------------|------------|------------|-------------|------|
| ATGGAGGAGC | CGCAGTCAGA | TCCTAGCGTC | GAGCCCCCTC | 40 |
| TGAGTCAGGA | AACATTTTCA | GACCTATGGA | AACTACTTCC | 80 |
| TGAAAACAAC | GTTCTGTCCC | CCTTGCCGTC | CCAAGCAATG | 120 |
| GATGATTGTA | TGCTGTCCCC | GGACGATATT | GAACAATGGT | 160 |
| TCACTGAAGA | CCCAGGTCCA | GATGAAGCTC | CCAGAATGCC | 200 |
| AGAGGCTGCT | CCCCGCGTGG | CCCCTGCACC | AGCAGCTCCT | 240 |
| ACACCGGCGG | CCCCTGCACC | AGCCCCCTCC | TGGCCCCCTGT | 280 |
| CATCTTCTGT | CCCTTCCCAG | AAAACCTACC | AGGGCAGCTA | 320 |
| CGGTTTCCGT | CTGGGCTTCT | TGCATTCTGG | GACAGCCAAG | 360 |
| TCTGCCACCA | TGGCCTACTC | CCCTGCGTCT | GTGACTTGCA | 400 |
| CGTACTCCCC | TGCCCTCAAC | AAGATGTTTT | GCCAAC TGGC | 440 |
| CAAGACCTGC | CCTGTGCAGC | TGTGGGTTGA | TTCCACACCC | 480 |
| CCGCCCCGCA | CCCGCGTCCG | CGCCATGGCC | ATCTACAAGC | 520 |
| AGTCACAGCA | CATGACGGAG | GTTGTGAGGC | GCTGCCCCCA | 560 |
| CCATGAGCGC | TGCTCAGATA | GCGATGGTCT | GGCCCCTCCT | 600 |
| CAGCGTCTTA | TCCGAGTGGA | AGGAAATTG | CGTGTGGAGT | 640 |
| ATTTGGATGA | CAGAAACACT | TTTCGACATA | GTGTGGTGGT | 680 |
| GCCCTATGAG | CCGCCTGAGG | TTGGCTCTGA | CTGTACCACC | 720 |
| ATCCACTACA | ACTACATGTG | TAACAGTTCC | TGCATGGGCG | 760 |
| GCATGAACCG | GAGGCCCATC | CTCACCATCA | TCACACTGGA | 800 |
| AGACTCCAGT | GGTAATCTAC | TGGGACGGAA | CAGCTTTGAG | 840 |
| GTGCGTGTTT | GTGCCTGTCC | TGGGAGAGAC | CGGCGCACAG | 880 |
| AGGAAGAGAA | TCTCCGCAAG | AAAGGGGAGC | CTCACCACGA | 920 |
| GCTGCCCCCA | GGGAGCACTA | AGCGAGCACT | GCCCAACAAC | 960 |
| ACCAGCTCCT | CTCCCCAGCC | AAAGAAGAAA | CCACTGGATG | 1000 |
| GAGAATATTT | CACCCTTCAG | ATCCGTGGGC | GTGAGCGCTT | 1040 |
| CGAGATGTTT | TTTGGTGACC | TACCTCTTCG | GAATTGCCGA | 1080 |
| GTCTTCCGAG | AGCTGAATGA | GGCCTTGGA | CTCAAGGATG | 1120 |
| CCCAGGCTGG | GAAGGAGCCA | GGGGGGAGCA | GGGCTCACTC | 1160 |
| CAGCCACCTG | AAGTCCAAAA | AGGGTCAGTC | TACCTCCCGC | 1200 |
| CATAAAAAAC | TCATGTTCAA | GACAGAAGGG | CCTGACTCAG | 1240 |
| AC | | | | 1242 |

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 608 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:26:

| | | | | |
|------------|------------|------------|------------|-----|
| CTCGGGCCGC | GTTGCTGGCG | TTTTTCCATA | GGCTCCGCCC | 40 |
| CCCTGACGAG | CATCACAAAA | ATCGACGCTC | AAGTCAGAGG | 80 |
| TGGCGAAACC | CGACAGGACT | ATAAAGATAC | CAGGCGTTTC | 120 |
| CCCCTGGAAG | CTCCCTCGTG | CGCTCTCCTG | TTCCGACCCT | 160 |
| GCCGCTTACC | GGATACCTGT | CCGCCTTTCT | CCCTTCGGGA | 200 |
| AGCGTGGCGC | TTTCTCAATG | CTCACGCTGT | AGGTATCTCA | 240 |
| GTTCGGTGTA | GGTCGTTTCG | TCCAAGCTGG | GCTGTGTGCA | 280 |
| CGAACCCCCC | GTTCAGCCCC | ACCGCTGCGC | CTTATCCGGT | 320 |
| AACTATCGTC | TTGAGTCCAA | CCCGGTAAGA | CACGACTTAT | 360 |
| CGCCACTGGC | AGCAGCCACT | GGTAACAGGA | TTAGCAGAGC | 400 |
| GAGGTATGTA | GGCGGTGCTA | CAGAGTTCTT | GAAGTGGTGG | 440 |
| CCTAACTACG | GCTACACTAG | AAGGACAGTA | TTTGGTATCT | 480 |
| GCGCTCTGCT | GAAGCCAGTT | ACCTTCGGAA | AAAGAGTTGG | 520 |
| TAGCTCTTGA | TCCGGCAAAC | AAACCACCGC | TGGTAGCGGT | 560 |
| GGTTTTTTTG | TTTGCAAGCA | GCAGATTACG | CGCAGAAAAA | 600 |
| AAGGATCT | | | | 608 |

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1547 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:27:

| | | | | |
|------------|------------|------------|------------|-----|
| GGTACCTGCC | ACCATGGCGC | GGATTCTTTA | TCACTGATAA | 40 |
| GTTGGTGGAC | ATATTATGTT | TATCAGTGAT | AAAGTGTCAA | 80 |
| GCATGACAAA | GTTGCAGCCG | AATACAGTGA | TCCGTGCCGG | 120 |
| CCCTGGACTG | TTGAACGAGG | TCGGCGTAGA | CGGTCTGACG | 160 |
| ACACGCAAAC | TGGCGGAACG | GTTGGGGGTG | CAGCAGCCGG | 200 |
| CGCTTTACTG | GCACTTCAGG | AACAAGCGGG | CGCCTTAAGG | 240 |
| GCCATATGGT | GAGTGGATGC | CTTGACCCCA | GGCGGGGATG | 280 |
| GGGGAGACCT | GTAGTCAGAG | CCCCCGGGCA | GCACAGGCCA | 320 |
| ATGCCCCGTC | TTCCCCGTGA | GGATGAGTAG | TGAGTGCCTC | 360 |
| TCCTGGCCCT | GGAAGTTGCC | ACTCCAGTGC | CCACCAGCCT | 400 |
| TGTCCTAATA | AAATTAAGTT | GCATCATTTT | GTCTGACTAG | 440 |
| GTGTCCTCTA | TAATATTATA | AGCTTGATAT | CGAATTCCTT | 480 |
| CGGACTTTTG | AAAGTGATGG | TGGTGGGGGA | AGGATTCGAA | 520 |
| CCTTCGAAGT | CGATGACGGC | AGATTTAGAG | TCTGCTCCCT | 560 |
| TTGGCCGCTC | GGGAACCCCA | CCACGGGTAA | TGCTTTTACT | 600 |
| GGCCTGCTCC | CTTATCGGGA | AGCGGGGCGC | ATCATATCAA | 640 |
| ATGACGCGCC | GCTGTAAAGT | GTTACGTTGA | GAAAGAATTC | 680 |
| CTGCAGCCCC | CCGCGTTGCT | GGCGTTTTTC | CATAGGCTCC | 720 |

| | | | | |
|------------|------------|------------|------------|------|
| GCCCCCTGA | CGAGCATCAC | AAAAATCGAC | GCTCAAGTCA | 760 |
| GAGGTGGCGA | AACCCGACAG | GACTATAAAG | ATACCAGGCG | 800 |
| TTTCCCCCTG | GAAGCTCCCT | CGTGCCTCT | CCTGTTCCGA | 840 |
| CCCTGCCGCT | TACCGGATAC | CTGTCCGCCT | TTCTCCCTTC | 880 |
| GGGAAGCGTG | GCGCTTTCTC | AATGCTCACG | CTGTAGGTAT | 920 |
| CTCAGTTCGG | TGTAGGTCGT | TCGCTCCAAG | CTGGGCTGTG | 960 |
| TGCACGAACC | CCCCGTTTCT | CCCGACCGCT | GCGCCTTATC | 1000 |
| CGGTAACAT | CGTCTTGAGT | CCAACCCGGT | AAGACACGAC | 1040 |
| TTATCGCCAC | TGGCAGCAGC | CACTGGTAAC | AGGATTAGCA | 1080 |
| GAGCGAGGTA | TGTAGGCGGT | GCTACAGAGT | TCTTGAAGTG | 1120 |
| GTGGCCTAAC | TACGGCTACA | CTAGAAGGAC | AGTATTTGGT | 1160 |
| ATCTGCGCTC | TGCTGAAGCC | AGTTACCTTC | GGAAAAAGAG | 1200 |
| TTGGTAGCTC | TTGATCCGGC | AAACAAACCA | CCGCTGGTAG | 1240 |
| CGGTGGTTTT | TTTGTTTGCA | AGCAGCAGAT | TACGCGCAGA | 1280 |
| AAAAAAGGAT | CTGGGGGATC | CGGAGAGCTC | ACTCTAGATG | 1320 |
| AGAGAGCAGT | GAGGGAGAGA | CAGAGACTCG | AATTTCCGGA | 1360 |
| GCTATTTTCT | TTTTCTTTTC | CGTTTTGTGC | AATTTCACTT | 1400 |
| ATGATACCGG | CCAATGCTTG | GTTGCTATTT | TGGAAACTCC | 1440 |
| CCTTAGGGGA | TGCCCCCTCA | CTGGCCCTAT | AAAGGGCCAG | 1480 |
| CCTGAGCTGC | AGAGGATTCC | TGCAGAGGAT | CAAGACAGCA | 1520 |
| CGTGGACCTC | GCACAGCCTC | TCCCACA | | 1547 |

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1807 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

| | | | | |
|------------|-------------|------------|-------------|-----|
| GGTACCTGCC | ACCATGGCGC | GGATTCTTTA | TCACTGATAA | 40 |
| GTTGGTGGAC | ATATTATGTT | TATCAGTGAT | AAAGTGTCAA | 80 |
| GCATGACAAA | GTTGCAGCCG | AATACAGTGA | TCCGTGCCGG | 120 |
| CCCTGGACTG | TTGAACGAGG | TCGGCGTAGA | CGGTCTGACG | 160 |
| ACACGCAAAC | TGGCGGAACG | GTTGGGGGTG | CAGCAGCCGG | 200 |
| CGCTTTACTG | GCACTTCAGG | AACAAGCGGG | CGCCTTAAGG | 240 |
| GCCATATGGT | GAGTGGATGC | CTTGACCCCA | GGCGGGGATG | 280 |
| GGGGAGACCT | GTAGTCAGAG | CCCCCGGGCA | GCACAGGCCA | 320 |
| ATGCCCCGTC | TTCCCCCTGCA | GGATGAGTAG | TGAGTGCCTC | 360 |
| TCCTGGCCCT | GGAAGTTGCC | ACTCCAGTGC | CCACCAGCCT | 400 |
| TGTCTTAATA | AAATTAAGTT | GCATCATTTT | GTCTGACTAG | 440 |
| GTGTCCTCTA | TAATATTATA | AGCTTGATAT | CGAATTCTTT | 480 |
| CGGACTTTTG | AAAGTGATGG | TGGTGGGGGA | AGGATTTCGAA | 520 |
| CCTTCGAAGT | CGATGACGGC | AGATTTAGAG | TCTGCTCCCT | 560 |
| TTGGCCGCTC | GGGAACCCCA | CCACGGGTAA | TGCTTTTACT | 600 |
| GGCCTGCTCC | CTTATCGGGA | AGCGGGGCGC | ATCATATCAA | 640 |
| ATGACGCGCC | GCTGTAAAGT | GTTACGTTGA | GAAAGAATTC | 680 |

| | | | | |
|-------------|------------|------------|------------|------|
| CTGCAGCCCCG | CCGCGTTGCT | GGCGTTTTTC | CATAGGCTCC | 720 |
| GGCCCCCTGA | CGAGCATCAC | AAAAATCGAC | GCTCAAGTCA | 760 |
| GAGGTGGCGA | AACCCGACAG | GACTATAAAG | ATACCAGGCG | 800 |
| TTTCCCCCTG | GAAGCTCCCT | CGTGCGCTCT | CCTGTTCCGA | 840 |
| CCCTGCCGCT | TACCGGATAC | CTGTCCGCCT | TTCTCCCTTC | 880 |
| GGGAAGCGTG | GCGCTTTCTC | AATGCTCACG | CTGTAGGTAT | 920 |
| CTCAGTTCGG | TGTAGGTCGT | TCGCTCCAAG | CTGGGCTGTG | 960 |
| TGCACGAACC | CCCCGTTTCA | CCCGACCGCT | GCGCCTTATC | 1000 |
| CGGTAACAT | CGTCTTGAGT | CCAACCCGGT | AAGACACGAC | 1040 |
| TTATCGCCAC | TGGCAGCAGC | CACTGGTAAC | AGGATTAGCA | 1080 |
| GAGCGAGGTA | TGTAGGCGGT | GCTACAGAGT | TCTTGAAGTG | 1120 |
| GTGGCCTAAC | TACGGCTACA | CTAGAAGGAC | AGTATTGTTG | 1160 |
| ATCTGCGCTC | TGCTGAAGCC | AGTTACCTTC | GGAAAAAGAG | 1200 |
| TTGGTAGCTC | TTGATCCGGC | AAACAAACCA | CCGCTGGTAG | 1240 |
| CGGTGGTTTT | TTTGTTTGCA | AGCAGCAGAT | TACGCGCAGA | 1280 |
| AAAAAAGGAT | CTGGGGGATC | CGGAGAGCTC | CCAACGCGTT | 1320 |
| GGATGCATGG | ATGAGGGAAA | GGAGGTAAGA | TCTGTAATGA | 1360 |
| ATAAGCAGGA | ACTTTGAAGA | CTCAGTGAAT | CAGTGAGTAA | 1400 |
| TAAAGACTCA | GTGACTTCTG | ATCCTGTCCT | AACTGCCACT | 1440 |
| CCTTGTTGTC | CCAAGAAAGC | GGCTTCCTGC | TCTCTGAGGA | 1480 |
| GGACCCCTTC | CCTGGAAGGT | AAAATAAGG | ATGTCAGCAG | 1520 |
| AGAAATTTTT | CCACCATTGG | TGCTTGGTCA | AAGAGGAAAC | 1560 |
| TGATGAGCTC | ACTCTAGATG | AGAGAGCAGT | GAGGGAGAGA | 1600 |
| CAGAGACTCG | AATTTCCGGA | GCTATTTTCA | TTTTCTTTTC | 1640 |
| CGTTTTGTGC | AATTTCACTT | ATGATAACCG | CCAATGCTTG | 1680 |
| GTTGCTATTT | TGGAAACTCC | CCTTAGGGGA | TGCCCTCAA | 1720 |
| CTGGCCCTAT | AAAGGGCCAG | CCTGAGCTGC | AGAGGATTCC | 1760 |
| TGCAGAGGAT | CAAGACAGCA | CGTGGACCTC | GCACAGCCTC | 1800 |
| TCCCACA | | | | 1807 |

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:29:

| | | | | |
|------------|------------|------------|------------|-----|
| GGTACCTGCC | ACCATGGCGA | AGGGCGAGGA | ACTGTTCACT | 40 |
| GGCGTGGTCC | CAATTCTCGT | GGAAGTGGAT | GGCGATGTGA | 80 |
| ATGGGCACAA | ATTTTCTGTC | AGCGGAGAGG | GTGAAGGTGA | 120 |
| TGCCACATAC | GGAAAGCTCA | CCCTGAAATT | CATCTGCACC | 160 |
| ACTGGAAAGC | TCCCTGTGCC | ATGGCCAACA | CTGGTCACTA | 200 |
| CCTTCACCTA | TGGCGTGCAG | TGCTTTTCCA | GATACCCAGA | 240 |
| CCATATGAAG | CAGCATGACT | TTTTCAAGAG | CGCCATGCCC | 280 |

| | | | | |
|-------------|------------|------------|------------|------|
| GAGGGCTATG | TGCAGGAGAG | AACCATCTTT | TTCAAAGATG | 320 |
| ACGGGAACTA | CAAGACCCGC | GCTGAAGTCA | AGTTCGAAGG | 360 |
| TGACACCCTG | GTGAATAGAA | TCGAGTTGAA | GGGCATTGAC | 400 |
| TTTAAGGAAG | ATGGAAACAT | TCTCGGCCAC | AAGCTGGAAT | 440 |
| ACAACTATAA | CTCCACAAT | GTGTACATCA | TGGCCGACAA | 480 |
| GCAAAAGAAT | GGCATCAAGG | TCAACTTCAA | GATCAGACAC | 520 |
| AACATTGAGG | ATGGATCCGT | GCAGCTGGCC | GACCATTATC | 560 |
| AACAGAACAC | TCCAATCGGC | GACGGCCCTG | TGCTCCTCCC | 600 |
| AGACAACCAT | TACCTGTCCA | CCCAGTCTGC | CCTGTCTAAA | 640 |
| GATCCCAACG | AAAAGAGAGA | CCACATGGTC | CTGCTGGAGT | 680 |
| TTGTGACCGC | TGCTGGGATC | ACACATGGCA | TGGACGAGCT | 720 |
| GTACAAGTGA | GCGCCTTAAG | GGCCATATGG | TGAGTGGATG | 760 |
| CCTTGACCCC | AGGCGGGGAT | GGGGGAGACC | TGTAGTCAGA | 800 |
| GCCCCGGGCG | AGCACAGGCC | AATGCCCGTC | CTTCCCCTGC | 840 |
| AGGATGAGTA | GTGAGTGCCT | CTCCTGGCCC | TGGAAGTTGC | 880 |
| CACCTCCAGTG | CCCACCAGCC | TTGTCTTAAT | AAAATTAAGT | 920 |
| TGCATCATTT | TGTCTGACTA | GGTGTCTCT | ATAATATTAT | 960 |
| AAGCTTGATA | TCGAATTCTT | TCGGACTTTT | GAAAGTGATG | 1000 |
| GTGGTGGGGG | AAGGATTCGA | ACCTTCGAAG | TCGATGACGG | 1040 |
| CAGATTTAGA | GTCTGCTCCC | TTTGGCCGCT | CGGGAACCCC | 1080 |
| ACCACGGGTA | ATGCTTTTAC | TGGCCTGCTC | CCTTATCGGG | 1120 |
| AAGCGGGGCG | CATCATATCA | AATGACGCGC | CGCTGTAAAG | 1160 |
| TGTTACAGTG | AGAAAGAATT | CCTGCAGCCC | GCCGCGTTGC | 1200 |
| TGGCGTTTTT | CCATAGGCTC | CGCCCCCCTG | ACGAGCATCA | 1240 |
| CAAAAATCGA | CGCTCAAGTC | AGAGGTGGCG | AAACCCGACA | 1280 |
| GGACTATAAA | GATACCAGGC | GTTTCCCCCT | GGAAGCTCCC | 1320 |
| TCGTGCGCTC | TCCTGTTCCG | ACCCTGCCGC | TTACCGGATA | 1360 |
| CCTGTCCGCC | TTTCTCCCTT | CGGGAAGCGT | GGCGCTTTCT | 1400 |
| CAATGCTCAC | GCTGTAGGTA | TCTCAGTTTC | GTGTAGGTCG | 1440 |
| TTTCGTTCCAA | GCTGGGCTGT | GTGCACGAAC | CCCCCGTTCA | 1480 |
| GCCGACCCGC | TGCGCCTTAT | CCGGTAACTA | TCGTCTTGAG | 1520 |
| TCCAACCCGG | TAAGACACGA | CTTATCGCCA | CTGGCAGCAG | 1560 |
| CCACTGGTAA | CAGGATTAGC | AGAGCGAGGT | ATGTAGGCGG | 1600 |
| TGCTACAGAG | TTCTTGAAGT | GGTGGCCTAA | CTACGGCTAC | 1640 |
| ACTAGAAGGA | CAGTATTTGG | TATCTGCGCT | CTGCTGAAGC | 1680 |
| CAGTTACCTT | CGGAAAAAGA | GTTGGTAGCT | CCTGATCCGG | 1720 |
| CAAACAAACC | ACCGCTGGTA | GCGGTGGTTT | TTTTGTTTGC | 1760 |
| AAGCAGCAGA | TTACGCGCAG | AAAAAAAGGA | TCTGGGGGAT | 1800 |
| CCGGAGAGCT | CCCAACGCGT | TGGATGCATG | GATGAGGGAA | 1840 |
| AGGAGGTAAG | ATCTGTAATG | AATAAGCAGG | AACTTTGAAG | 1880 |
| ACTCAGTGAC | TCAGTGAGTA | ATAAAGACTC | AGTGACTTCT | 1920 |
| GATCCTGTCC | TAACTGCCAC | TCCTTGTTGT | CCCAAGAAAG | 1960 |
| CGGCTTCCTG | CTCTCTGAGG | AGGACCCCTT | CCCTGGAAGG | 2000 |
| TAAAACTAAG | GATGTCAGCA | GAGAAATTTT | TCCACCATTG | 2040 |
| GTGCTTGGTC | AAAGAGGAAA | CTGATGAGCT | CACTCTAGAT | 2080 |
| GAGAGAGCAG | TGAGGGAGAG | ACAGAGACTC | GAATTTCCGG | 2120 |
| AGCTATTTCA | GTTTTCTTTT | CCGTTTTGTG | CAATTTCACT | 2160 |
| TATGATACCG | GCCAATGCTT | GGTTGCTATT | TTGGAAACTC | 2200 |
| CCCTTAGGGG | ATGCCCTCA | ACTGGCCCTA | TAAAGGGCCA | 2240 |
| GCCTGAGCTG | CAGAGGATTC | CTGCAGAGGA | TCAAGACAGC | 2280 |
| ACGTGGACCT | CGCACAGCCT | CTCCCACA | | 2308 |